Hyperspectral Image Classification Using k-means Clustering

Sameer Ranjan

Department of Computer Science and Engineering

National Institute of Technology Rourkela

Rourkela-769 008, Odisha, India

Hyperspectral Image Classification using k-means Clustering

Thesis submitted in partial fulfillment

of the requirement for the degree of

Bachelor of Technology

in

Computer Science and Engineering

by

Sameer Ranjan

(111CS0483)

under the guidance of

Prof. Banshidhar Majhi

Department of Computer Science and Engineering

National Institute of Technology Rourkela

Rourkela-769 008, Odisha, India

May 2015

Department of Computer Science and Engineering

National Institute of Technology Rourkela

Rourkela-769 008

10th May 2015

Certificate

This is to certify that the work in the thesis entitled *Hyperspectral Image Classification using k-means Clustering* by *Sameer Ranjan* is a record of original research work carried out under my supervision and guidance in partial fulfillment of the requirement for the award of the degree of Bachelor of Technology in Computer Science and Engineering. Neither this thesis nor any part of it has been submitted for any degree or academic award elsewhere.

Prof. Banshidhar Majhi

Professor

Acknowledgment

I would like to express my gratitude and appreciation to all those who helped me to complete this thesis.

A special thanks to my project advisor, Prof. Banshidhar Majhi whose help, stimulating suggestions and encouragement helped me to coordinate my project. He provided his able guidance whenever I needed it. His profound insight has guided my thinking to improve the final product. My earnest gratefulness to him.

I am grateful to Prof. Ratnakar Dash for providing his valuable suggestion without which the project would have not been completed. I am also grateful to Rig Das for his support and ideas and always being there for me.

Finally,I am grateful to all of my friend for continuous motivation and encouragement.Last but not least to my family having faith in me and always supporting me.

Sameer Ranjan

Dedicated to my Parents

Abstract

Hyperspectral Image stores the reflectance of objects across the electromagnetic spectrum. Each object is identified by its spectral signature. Hyperspectral Sensors records these images from airborne devices. By processing these Images we can get various information about the land-form, seabed etc.

This thesis presents an efficient and accurate classification technique for Hyperspectral Images. The approach consists of three steps. Firstly, dimension reduction of the Hyperspectral Image using Principal Component Analysis. This is done in order to reduce the time complexity of the further process. Secondly, the reduced features are clustered by k-means clustering. Lastly, the clusters are individually trained by Support Vector Machine.

This scheme was tested with Pavia University Data-set taken by ROSIS sensor. Using the above scheme overall accuracy of 90.2% was achieved which is very promising in comparison to conventional Support Vector Machine classification which had an overall accuracy of 78.67% with the same data-set.

Contents

List of Figures

Chapter 1

Introduction

Over the last few years, there has been a remarkable increase in the number of remote sensing sensors on-board various satellite and aircraft platforms. Noticeable is the availability of data from hyperspectral sensors such as AVIRIS, HYDICE, HyMap and HYPERION. The hyperspectral data together with geographical information system (GIS) derived ancillary data form an exceptional spatial database for any scientific study related to Earth's environment. Thus, significant advances have been made in remote sensing data acquisition, storage and management capabilities. The availability of huge spatial databases brings in new challenges for the extraction of quality information. The sheer increase in the volume of data available has created the need for the development of new techniques that can automate extraction of useful information to the greatest degree. Moreover, these techniques need to be objective, reproducible, and feasible to implement within available resources.

1.1 Spectral Image

Spectral Images are three dimensional matrices which is the product of stacking several two dimensional(x-axis and y axis)images together. Each of these image stores the reflectance data of the land-form of a particular wavelength of Infrared Regions. These wavelength are have a contagious range(400-2500nm) with a step of some nanometers. The resultant third dimension is representing the spectral data representing discrete spectral bands. So each pixel in a Spectral image is represented by a vector which is also called spectral signature. Based on the gap of wavelength between two contagious band there are three types of Spectral Images.

1.1.1 Multispectral Image

The spectral resolution of this kind of band is (100-200)nm. Major setback to this image is, narrow spectral feature may not be easily discriminated and have the tendency to average out on spectral sampling or are masked by nearby strong feature, resulting in the loss in the information that can be readily extracted from it.

1.1.2 Hyperspectral Image

The spectral resolution of this kind of band is 5-10nm. In 1983 (Airborne Visible Infrared Imaging Spectrometer) AVIRIS was proposed. This sensor was first to acquire image with spectral resolution of 5-10nm. With this development there was an encouragement in the area of Hyperspectral Imaging.

Within these images we can easily identify and quantify the different types of landform by studying the molecular absorption and particle scattering and the spectral signature. Specific spectral signature can be associated with different chemical, biological,mineral processes going on the land form which gives us the opportunity to differentiate it. [1].

1.1.3 Ultraspectral Image

This image has a band gap of less than 5mm. It contain even more spectral bands than Hyperspectral Image. It is considered as future of Spectral Imaging. Various ultraspectral sensors are being manufactured to encourage the use of Ultraspectral Imaging.

1.2 Advantage of Hyperspectral Imaging

Most of the multispectral images stores the reflectance of the earth in the wide wavelength range, and in between the range no measurement of intensity is stored. In contrast, hyperspectral image stores the intensity information of the narrow wavelength bands, this is very desirable as every minute information about earth is getting stored. The spectrum measured by hyperspectral sensor is much like the spectrum measured in spectroscopy laboratory. So the detailed spectrum can provide every detailed information that a traditional multispectral spectrum cannot provide. [2].

1.3 Frequent Terms used in Hyperspectral Imaging

1.3.1 Spectral Resolution

As discussed above spectral resolution is the narrowest bandwidth over which reflectance is recorded or in simple words the minimum wavelength gap between two successive bands. This is a property of sensor and is a critical parameter for determining the detailing of the information measured in the image.

1.3.2 Spacial Resolution

The spacial resolution of a sensor refers to the distance between the nearest object that can be recorded in the unit of length. The detailing of the image in the terms of length is depending on spacial resolution of the sensor as it refers to the smallest object that can be detected. It depends primarily on their field of view of that moment.

1.3.3 Temporal Resolution

Temporal resolution refers to the frequency by which a sensor can obtain information. As repeated information about the landform can be helpful in monitoring the land form as, a slight change can be detected in two subsequent instances. The potential benefits of these repeated information also lies in the fact that they provide information about the landform upon a length of time and that change can help us study the behavior of the lanform over time.

1.3.4 Radiometric Resolution

It is defined by the sensitivity of the sensor to the difference in the strength of the electromagnetic radiation(EMR)signal and determines the smallest differences in the intensity of the signal that can be distinguished.

1.4 Hyperspectral Image Classification

Hyperspectral image classification is a tool for remote sensing which involves in analysis and interpretation of data acquired from sensors. the Hyperspectral classification works on the principle that the spectral signature for each object is unique and it is based on its ability to reflect light, which is dependent on the chemical, biological behavior of that element. This property is used by the researchers and they see the similarity between two object and they classify the object as one class. Hyperspectral image classification also attracts interest of the world in its applications in urban planning, agriculture and monitoring. With the help of hyperspectral classification we can generate thematic maps showing different classification classes in a pictorial way of the land cover. [3].

1.5 Structure of this Thesis

Chapter 2 contains the details of the existing work done in the area of Hyperspectral Image classification. Chapter 3 describes the feature extraction process i.e Principal Component Analysis. Chapter 4 describes k-means Clustering which is an important component of this project. Chapter 5 describes the Support Vector Machine which is a tool used to classify the Hyperspectral Image. Chapter 4 contain the detail description of the methods used in this project. Chapter 5 shows the results and simulation and the data obtained. Chapter 6 is a brief conclusion with a possibility of future work.

Chapter 2

Literature Survey

Craig Rodarmel and Jie Shan have explained the process of Principal Component Analysis. They have used HYDICE and AVIRIS dataset for this process. The content analysis shows that the information content decreases with increasing band number. Hence by using the first few principal components they are able to achieve 70% classification accuracy. This indicate that the use of principle component analysis as a feature extraction technique can yield a better classification of hyperspectral image [4].

Mark Richardson has provided technical details of Principal Component Analysis. He has used the Singular Value Decomposition in the process. The algorithm provided was very useful in implementing Principal Component Analysis in this project [5].

Aleix M. Martnez, Member and Avinash C. Kak have presented a comprehensive analysis between Principal Component Analysis and Linear Discriminant Analysis. It is evident that PCA though less powerful can outperform LDA in case of small dataset. They have also concluded that PCA is less sensitive to different data sets [6].

Kanungo *et al.* have presented an efficient and simple implementation of Lloyd's k-means clustering algorithm, also known as the filtering algorithm. They presented a data-sensitive analysis of the algorithm's running time, which shows that the algorithm runs faster as the separation between clusters increases [7].

Michael Steinbach *et al.* have provided an introduction on cluster analysis, and then focused on the challenges of clustering high dimensional data. They also presented a brief overview of some recent clustering techniques, including their own technique which uses the concept of an efficient high dimension clustering. [8].

Rig Das *et al.* have presented a hyperspectral image classification method based on Quadratic Fishers Discriminant Analysis (QFDA) and Multi-class Support Vector Machine (M-SVM). They have utilized QFDA for feature extraction and dimensionality reduction of AVIRIS hyperspectral image and have used M-SVM for classification tool. They have compared their suggested scheme with Principal Component Analysis (PCA) with SVM scheme. It was observed that the method proposed by them is superior in terms of classification accuracy. [9].

Camps-Valls presented a semi-supervised graph-based method for the classification of hyperspectral images. The method is designed to handle the special characteristics of hyperspectral images, namely, high-input dimension of pixels, low number of labeled samples, and spatial variability of the spectral signature. The presented semi-supervised-graph-based method is compared to state-of-the-art support vector machines in the classification of hyperspectral data [10].

Alberto Villa *et al.* have used Independent Component (IC) Discriminant Analysis (ICDA) for remote sensing classification. ICDA is a non parametric method for discriminant analysis based on the application of a Bayesian classification rule on a signal composed by ICs. They have produced a classification over all accuracy of 82.14% [11].

C. Shah has presented a summary of our ongoing research on the classification of hyperspectral images. He has experimented with both supervised as well as unsupervised algorithms. His research has provided a benchmark for new researchers [12].

Marta Rojas *et al.* have evaluated the performance of different processing chains resulting from combinations of modules for dimensionality reduction, feature extraction/ selection, image classification, and spatial post-processing. They have used support vector machine in the project to classify the hyperspectral data because it has the ability to have a good accuracy even the sample size is small. They tested their scheme on AVIRIS Indian Pine dataset. [13].

Quanquan Gu and Jiawei Han have proposed a Clustered Support Vector Machine (CSVM), which apply divide and conquer approach in clustering. they divided the data into several clusters and then each cluster was trained separately by the M-S.V.M technique. This idea is the backbone of this thesis [14].

Yukai Yao *et al.* have presented an algorithm that uses k-means clustering algorithm and a SVM classifier to train and test a sample. [15]

Chapter 3

Principal Component Analysis

The principal component analysis is based on eigenvalue decomposition of co-variance matrix. It is used as a feature extraction technique. In a hyperspectral image two subsequent band are very close, hence they often carry the same information. This means that the data is correlated. Principal Component analysis is a tool to not only extract features out of an image but also decrease the correlation by decreasing the dimension of the image.

3.1 The Method

Let us consider an Hyperspectral Image as shown in 3.1. An Image pixel vector is a vector of all values falling at the same $x - y$ coordinate.

$$X_i = [x_1, x_2, ..., x_N]^T \tag{3.1}$$

where N is number of Hyperspectral band. for image with m rows and n columns there will be $M = m*n$ such vectors i.e $i = 1, 2, ..., M$. The mean of all the image pixel vectors can be calculated as:

$$m = \frac{1}{M}\sum_{i=1}^{M}[x_1, x_2, ..., x_N]_i^T \tag{3.2}$$

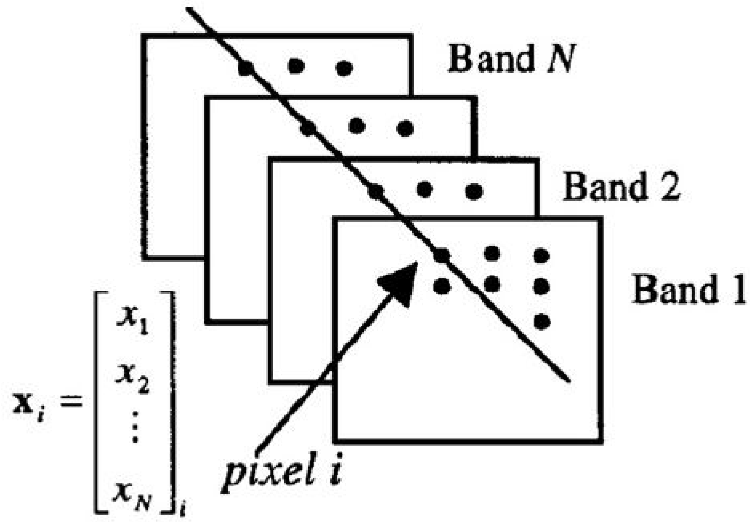

Figure 3.1: Pixel Vector in a Hyperspectral Image.

The covariance matrix can be calculated as:

$$C_X = \frac{1}{M}\sum_{i=1}^{M}(X_i - m)(X_i - m)^T \tag{3.3}$$

For eigenvalue decomposition of the covariance matrix, we can write:

$$C_X = ADA^T \tag{3.4}$$

where:

$$D = diag(\lambda_1, \lambda_2, ..., \lambda_N) \tag{3.5}$$

is the diagonal matrix composed of the eigenvalues $\lambda_1, \lambda_2, ..., \lambda_N$ of the covariance matrix C_X and A is the orthonormal matrix composed of the corresponding N dimension eigenvectors $a_k(k = 1, 2, ..., N)$ as follows:

$$A = (a_1, a_2, ..., a_N) \tag{3.6}$$

The linear transformation defined by:

$$Y_i = A^T X_i (i = 1, 2, ..., M) \tag{3.7}$$

is the PCA pixel vector, and all these pixel vectors form the PCA bands of the original images. Let the eignevectors be arranged in decreasing order of the eigen values so that $\lambda_1 >= \lambda_2 >= ... >= \lambda_N$, now the first K rows of the matrix A^T forms A_K^T and the transformation is modified as:

$$Y_i = A_K^T X_i (i = 1, 2, ..., M) \tag{3.8}$$

where K is the no. of Principal Components.Hence each image pixel vector is mapped using the above transformation. The new image with reduced band is obtained by placing each modified pixel vectors in the image at its pixel position [4].

3.2 Significance of Principal Component Analysis

The significance of PCA lies with the reduction of dimension that it performs of the image. The hyperpectral image has a spectral band of length 200 values. Processing this type of data is very cumbersome and time taking. The increased complexity can be a problem in the later subsequent steps. Hence the reduced dimension serves as an important element of PCA. It also decreases the correlation between the data.

3.3 Advantage of Principal Component Analysis

Including only first few principal components we can cover large percentage of total co-variance present in the result.

PCA is completely non-parametric i.e. any data set can be used as an input to get the output, requiring no parameters, and information about how the data was recorded.

It uses only second order statistics which is important for a large dataset in order to reduce the complexity of the computation.

Chapter 4

k-means Clustering

Clustering is the process of dividing a group of data into a number of smaller groups known as clusters. k-means clustering is one of the most popular and simple unsupervised learning algorithms that solve the clustering problem.

4.1 The Algorithm

The procedure follows a simple process in order to classify a given data-set in a number of clusters which is fixed a priori. The main idea of the algorithm is based upon defining k centroid for each clusters, from the given data-set. This selection is random and the results of clustering depends upon this selection, the best selection strategy is to place them as far as possible from each other. The next step is to pick each point in the data-set and assign it to the clusters whose centroid is nearest to the point. After this step an early grouping is done. Now the centoid for the clusters are recalculated from the grouping from previous step. Now after new centroid assigning again the points are reassigned as done previously. These two steps are repeated in loop. The loop terminates when no changes in the cetroid is there in two subsequent iterations. Finally, this algorithm has a objective function which has to be minimized, i.e. means square error which is calculated as:

$$mean\_square\_error = \sum_{i=1}^{k} \sum_{vector \in cluster_i} (vector - m_i)^T (vector - m_i) \tag{4.1}$$

where k is the number of clusters and m_i is the mean of the i^{th} cluster.

Algorithm 1: k-means

input : A Matrix X of size $m \times n$ and no of clusters k

output: k cluster and mean[k]

```
initialize k means by random k vectors from the input matrix X;
mean_sq_error ← 0;
new_mean_sq_error ← 0;
do
    mean_sq_error ← new_mean_sq_error;
    for vector ∈ X do
        calculate the index i such that vector has min euclidean distance from mean[i];
        add vector to cluster[i];
    end
    for i ← 1 to k do
        count ← 0;
        for vector ∈ cluster[i] do
            count = count + 1;
            mean[i] = mean[i] + vector;
        end
        mean[i] = mean[i]/count;
    end
    recalculate new_mean_sq_error;
while |mean_sq_error − new_mean_sq_error| ≠ 0;
```

4.2 Initializing Mean

Initialization of means is the step which decides that what will be the accuracy of classification and plays a very important role in clustering The most common method for initialization of means are Forgy Partition and Random Partition. In Forgy method k observations are randomly chosen from the data-set and are initialized as means. The random Partition method first assigns a cluster to each observation and then proceeds to the update step. The Forgy method of initialization is preferable.

4.3 Advantage of k-means clustering

The major advantage of this process is, this method is robust,efficient and easy to understand. If variables are huge, then k-means most of the times computationally faster than other clustering method, if we keep k small. k-means produce tighter clusters than other clustering method, especially if the clusters are globular.

Chapter 5

Support Vector Machine

5.1 Derivation of Support Vector Machine

A SVM is a binary classifier which creates a hyperplane which act as a decision boundary in a multi-dimensional space having training vectors. The hyperplane is decided using a subset of this training vectors which are also called support vectors. A nonlinear ψ-function is chosen that maps the input to higher dimensional space. We assume that each vector x_k has been transformed to $y_k = \psi(x_k)$. For each n patterns, $k = 1, 2, ..., n$, $Z_k = \pm 1$, where the vector may be in ω_1 or ω_2 respectively. A linear discriminant in an augmented y space is:

$$g(y) = a_t y \tag{5.1}$$

Thus, a separating hyperplane ensures:

$$Z_k G(y_k) \geq 1, \qquad where \quad k = 1, 2, ..., n \tag{5.2}$$

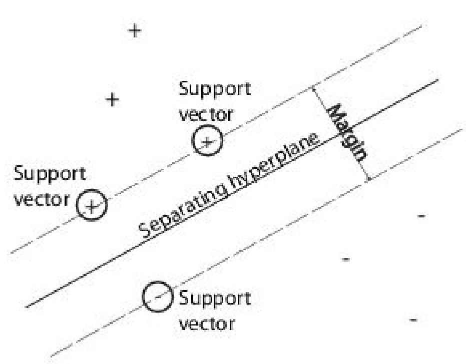

Figure 5.1: Support Vector Machine separating data by a Hyperplane.

SVM find the separating hyperplane with largest margin. The generalization of the classifier depends on the margin, larger the margin, the better the classifier. The distance between any hyperplane to a pattern y is given by $\frac{|g(y)|}{||a||}$, and let us assume a positive margin b exists, then Eq.5.2 implies:

$$\frac{Z_k g(y_k)}{||a||} \geq b, \qquad where \quad k = 1, 2, ..., n \tag{5.3}$$

The goal is to minimize $||a||$ so a function is defined as:

$$L(a, a) = \frac{1}{2}||a||^2 - \sum_{k=1}^{n} a_k[Z_k a^t y_k - 1] \tag{5.4}$$

and try to minimize $L(.)$ with respect to the weight vector a, and maximize it with respect to the undermined multipliers $\alpha_k \geq 0$. Using Kuhn-Tucker [?] construction, we can maximize

$$L(a) = \sum_{k=1}^{n} a_i - \frac{1}{2}\sum_{k=1}^{n} a_k a_j z_k z_j y_j^t y_k \tag{5.5}$$

subjected to constraints:

$$\sum_{k=1}^{n} z_k a_k = 0, \qquad a_k \geq 0, \qquad k = 1, 2, ..., n \tag{5.6}$$

5.2 Multi Class problem

Support Vector Machine is basically a binary classifier can only classify samples having two samples as shown in the above section.It can not be directly used to classify data of more than one class as it can define only one hyper-plane which can separate data into two group. However the samples in the real world can contain classes of more than two type, and Hyperspectral Image contains many classes like corn field, paved land, grass field etc. So the pure Support Vector Machine seems incompatible with this kind of problem. To overcome this problem we come up with a different approach. We train our sample for each class assigning that class sample as 1 and every other class as 0, if we are having k classes then we have to train and store the module for $k - 1$ classes. We have to store all the module because will collectively decide the class of the test set. While testing a data point it should be tested against all the module unless a match is found, for example if we have a data point a we will test it with first module and check if a belongs to the first class, if not then second module is tested, if this testing goes on to $k - 1$ module and result is negative then the point is assigned to $k^t h$ class. Hence following this algorithm we can classify data having more than two class

Chapter 6

Proposed Method

6.1 The Dataset

These are some Hyperspectral data capture by hyperspectral sensors and are easily available on internet.

6.1.1 Indian Pine

This scene was gathered by AVIRIS sensor over the Indian Pines test site in North-western Indiana and consists of 145×145 pixels and 224 spectral reflectance bands in the wavelength range $0.4 - 2.5 \times 10^6$ meters. This scene is a subset of a larger one. The Indian Pines scene contains two-thirds agriculture, and one-third forest or other natural perennial vegetation.

6.1.2 Pavia University

These are two scenes acquired by the ROSIS sensor during a flight campaign over Pavia, northern Italy. The number of spectral bands is 102 for Pavia Center and 103 for Pavia University. Pavia Center is a 1096×1096 pixels image, and Pavia University is 610×610 pixels, but some of the samples in both images contain no information and have to be discarded before the analysis.

6.1.3 Botswana

The NASA EO-1 satellite acquired a sequence of data over the Okavango Delta, Botswana in 2001-2004. The Hyperion sensor on EO-1 acquires data at 30 m pixel resolution over a 7.7 km strip in 242 bands covering the 400-2500 nm portion of the spectrum in 10 nm windows.

6.2 Proposed Method

The block diagram given in Figure 6.1 depicts the working of the Proposed Method. The first step in the method include Feature Extraction by Principal Component Analysis. We provide a Hyperspectral Image and apply PCA to reduce the number of bands. This step reduce the dimension of the image which reduces the complexity of the further method. The next step includes clustering of the image pixel vectors from the previous step. The clustering of the image pixel vector is according to the Algorithm 1 provided in chapter 4. The output of this algorithm is the group of clusters and and the array of means of these clusters. Now the training of these clusters are according to the Algorithm 2 and this Algorithm stores the SVM module for each Clusters.

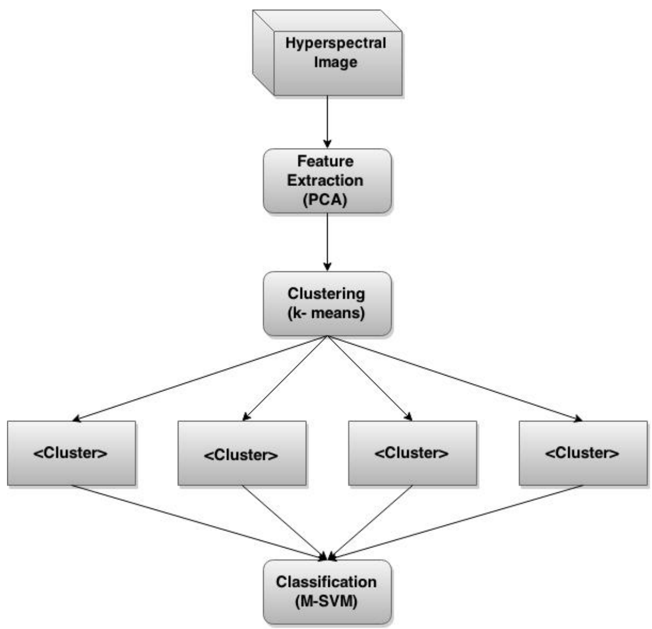

Figure 6.1: Block Diagram of the proposed method for classification.

After the training of the train-set the test-set is tested against the stored modules according to the Algorithm 3. The efficiency is calculated as:

$$OverallAccuracy(O.A.) = \frac{n(successful\_classification)}{n(test\_set)} \tag{6.1}$$

$$AverageAccuracy(A.A.) = \frac{1}{N}\sum_{i=1}^{N} E[i] \tag{6.2}$$

where E[i] is the accuracy for i^{th} cluster.

Algorithm 2: Training accordind to proposed method

input : A group of cluster C[k] where k is the no of cluster. A group of class label L[k]

output: A array of model Mod[k], module for each cluster

```
for i ← 1 to k do
    Mod[i]=msvm_train(C[i],L[i]);
end
```

Algorithm 3: Testing according to Proposed Method

input : Model Mod[k] from Algorithm 2, mean M[k] of the cluster and array of test-set T[M]

output: An array of result class for test set R[M]

```
for i ← 1 to M do
    index ← 0;
    dist ← eucled_dist(T[i], M[1]);
    for j ← 1 to k do
        temp_dist ← eucled_dist(T[i], M[j]);
        if temp_dist < dist then
            dist ← temp_dist;
            index ← j;
        end
    end
    R[i] ← msvm_test(Mod[index], T[i]);
end
```

Chapter 7

Results and Simulation

All the simulations for this project were done in MATLAB 2012a. Figure7.1 shows the Principal Components for the Indian Pine Dataset with number of component equal to 10. The original dataset contained 220 band which was reduced to 10. We can see that the maximum information component is present in first few bands.

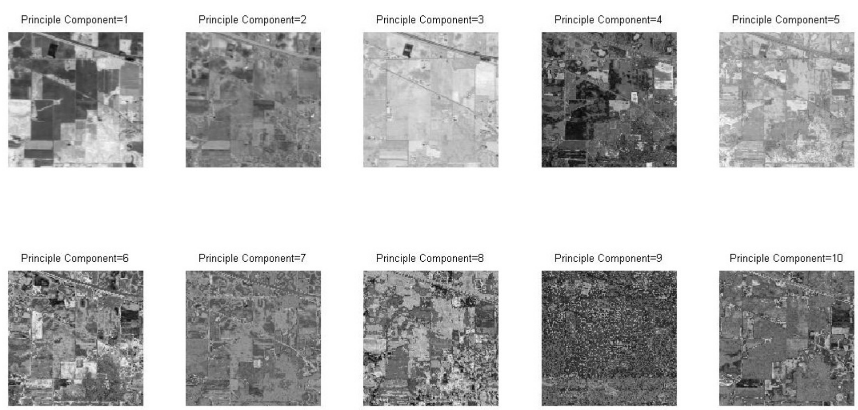

Figure 7.1: Principal Components after Principal Component Analysis on Indian Pine Data for $n = 10$

Figure 7.2 shows the simulation of clustering on the reduced Indian Pine Data which three types of distance in the algorithm viz. Manhattan Distance, Euclidean Distance, and Chessboard Distance. While Clustering the mean square error was calculated for each iteration and was plotted. It is evident from the plot that for each distance the mean square error drastically decrease in the first few iteration and then stabilizes at some constant value. It is also evident that Euclidean Distance provide better result than other two distance.

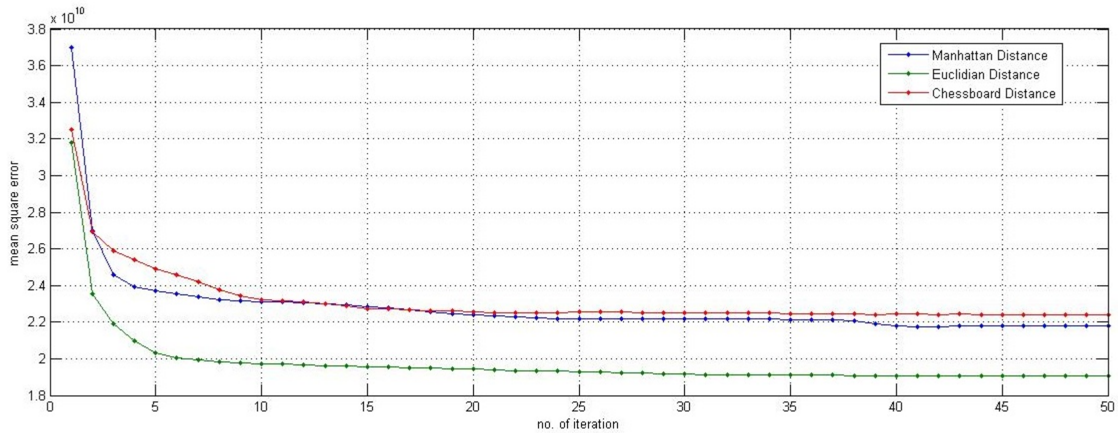

Figure 7.2: Plot for mean square error in clustering for different distances

Figure 7.3 shows the class composition of the clusters obtained by the clustering of the Indian Pine data. These are the histogram of each clusters which is this case is 16. Each histogram shows that how many classes is there for each class as the class no is represented on the x-axis. We can see that a single cluster is containing more that one class with one or two dominating class.

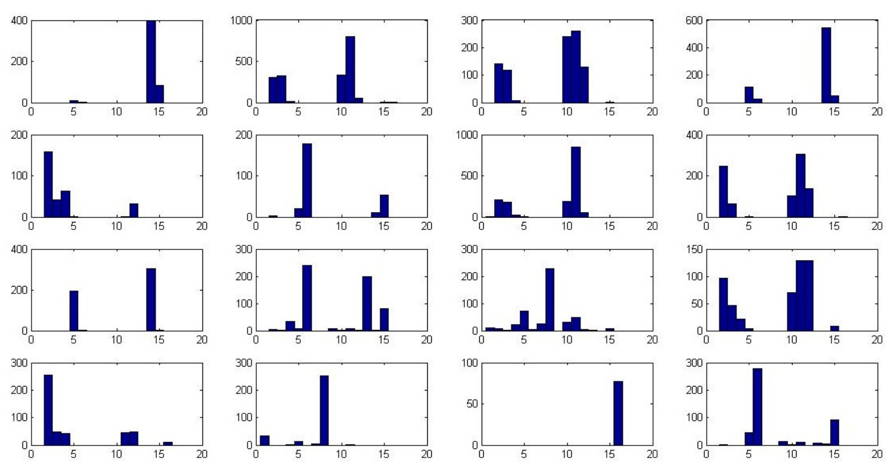

Figure 7.3: Class composition of cluster obtained after clustering Indian Pine Image

Table 7.1: Classification result for different Dataset.

Classification Accuracy						
	Pavia Univ.		Indian Pine		Botswana	
	S.V.M.	Proposed	S.V.M.	Proposed	S.V.M.	Proposed
Class 1	98.70%	85.50%	75.00%	57.14%	99.99%	99.99%
Class 2	96.86%	93.73%	89.24%	85.87%	96.29%	99.99%
Class 3	87.18%	81.69%	61.38%	70.81%	95.31%	96.97%
Class 4	78.27%	95.59%	06.78%	47.88%	93.44%	94.29%
Class 5	99.26%	99.75%	76.58%	92.80%	80.68%	77.92%
Class 6	56.78%	91.75%	73.13%	92.07%	77.77%	60.97%
Class 7	00.00%	74.18%	80.00%	88.88%	92.68%	98.78%
Class 8	19.91%	79.32%	61.74%	96.68%	87.69%	96.61%
Class 9	00.37%	99.25%	25.00%	33.33%	50.50%	79.38%
Class 10	-	-	63.42%	72.82%	75.00%	82.45%
Class 11	-	-	10.27%	65.61%	60.71%	88.46%
Class 12	-	-	13.40%	55.62%	81.13%	87.03%
Class 13	-	-	58.82%	96.42%	60.00%	90.54%
Class 14	-	-	79.41%	91.37%	72.41%	99.99%
Class 15	-	-	20.90%	75.85%	-	-
Class 16	-	-	00.00%	72.00%	-	-
Overall Accuracy	78.78%	90.22%	51.42%	77.15%	78.91%	88.19%
Average Accuracy	59.71%	88.97%	49.69%	74.70%	80.26%	89.53%
Execution Time	12003.21 sec	606.42 sec	1691.20 sec	168.96 sec	81.96 sec	8.16 sec

Table 7.1 shows the classification result of the three Dataset used in the project. For each Dataset we present the class wise classification efficiency for SVM classification and classification by proposed method. At last overall efficiency and average efficiency is shown which is calculated as:

$$OverallAccuracy(O.A.) = \frac{n(successful\_classification)}{n(test\_set)} \tag{7.1}$$

$$AverageAccuracy(A.A.) = \frac{1}{N}\sum_{i=1}^{N} E[i] \tag{7.2}$$

where E[i] is the accuracy for i^{th} cluster.

Figure 7.4 shows the overall accuracy and execution time for variable no of clusters. It can be visualized that the proposed method with 1 cluster is same as the traditional PCA and M-SVM method. We can see that if we go on increasing the number of clusters the overall accuracy increases and the execution time decreases at first then it becomes more or less constant. This graph is plotted by varying the number of clusters from 1 to 30 with a step of 2.

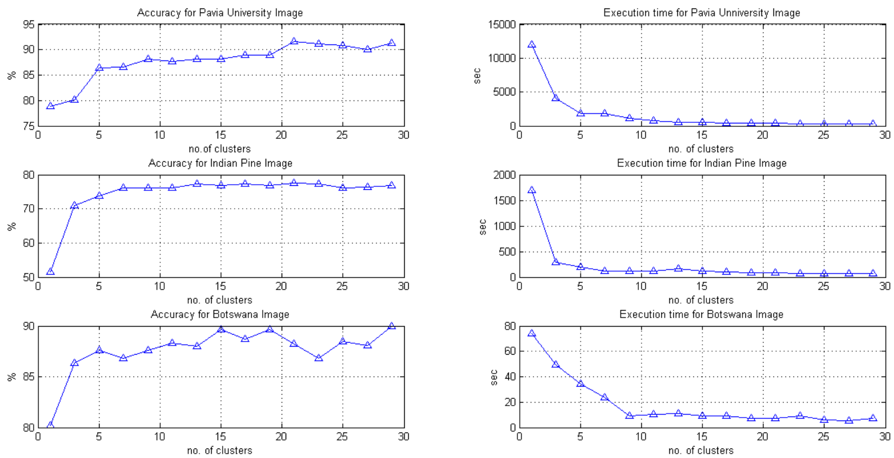

Figure 7.4: Plot for accuracy and execution time for variable number of clusters

The following Figure 7.5, Figure 7.6, and Figure 7.7 shows the classification result in a different manner. Figure(a) in all the figure shows the ground truth of the original dataset. Ground Truth of any hyperspectral image shows the class composition in a pictorial manner. All the pixels which represents the land cover are given values as their class number. Each class number is assigned with a rgb color and the image for the ground truth is obtained. Figure(b) shows the classification of the whole scene using the proposed method. Each pixel is classified as a class and the color used for the classes previously. The pixel which was unclassi

ed in the ground truth is also classified with the assumed accuracy of the Overall Accuracy which was obtained in the table 7.1 (c) shows the a classification of the whole scene by SVM. It is evident that the proposed method is producing more promising results.

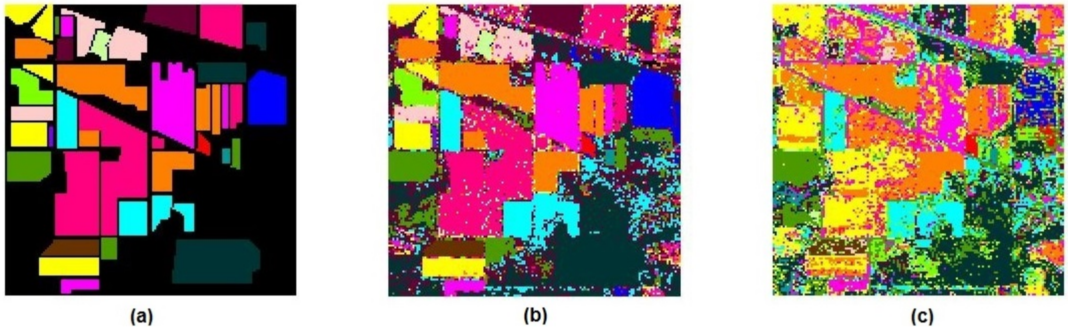

Figure 7.5: Classification result Pavia University Data a)Ground Truth of the Data b)Classification by proposed method c)Classification by SVM.

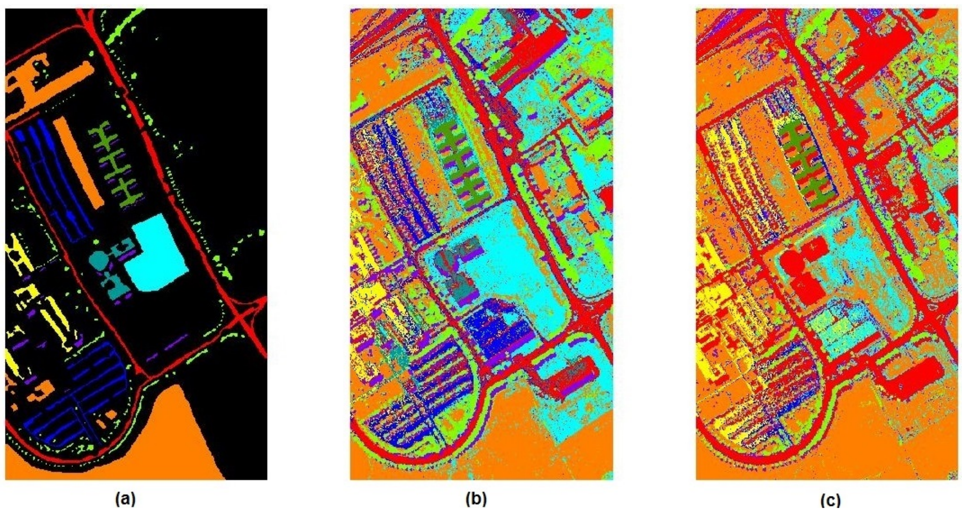

Figure 7.6: Classification result Pavia University Data a)Ground Truth of the Data b)Classification by proposed method c)Classification by SVM.

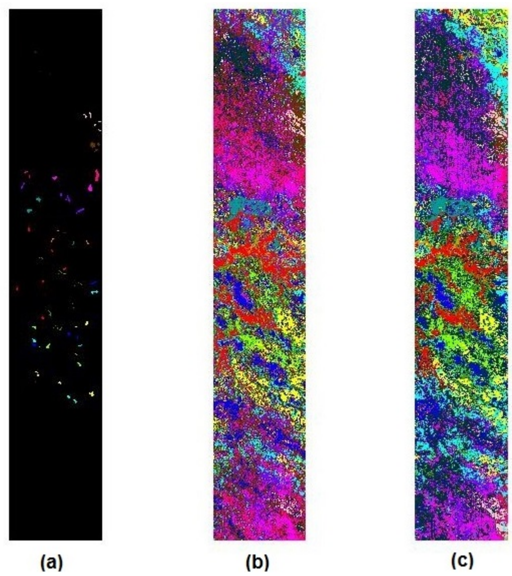

Figure 7.7: Classification result Pavia University Data a)Ground Truth of the Data b)Classification by proposed method c)Classification by SVM.

Chapter 8

Conclusion and Future Work

In this thesis, a method for classification of Hyperspectral Image is presented. This method presents dimension reduction by Principal Component Analysis, clustering of the reduced data by k-means clustering, and classification using Support Vector Machine. This scheme was tested with Pavia University Data-set taken by ROSIS sensor. Using the above scheme overall accuracy of 90.2% was achieved which is very promising in comparison to conventional Support Vector Machine classification which had an overall accuracy of 78.67% with the same data-set. Other Datasets are also tested to get promising result eg. Indian Pine, Botswana.

Future Work can include clustering of the clusters having many classes. This can improve the accuracy, because the clusters which have more than three or four classes are correlated and need a further clustering. Only after further clustering these classes can separate these data and the resulting child clusters will contain less no. of classes.

Bibliography

[1] P. Varshney and M. Arora, *Advanced Image Processing Techniques for Remotely Sensed Hyperspectral Data*. Springer, 2004.

[2] P. Shippert, "Why use hyperspectral imagery?," *Photogrammetric engineering and remote sensing*, vol. 70, no. 4, pp. 377–396, 2004.

[3] R. Ablin and C. H. Sulochana, "A survey of hyperspectral image classification in remote sensing," *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 2, pp. 2986–3003.

[4] J. S. Craig Rodarmel, "Principal component analysis for hyperspectral image classification," *Surveying and Land Information Systems*, vol. 62, no. 2, pp. 155–123, 2002.

[5] M. Richardson, "Principal component analysis," *URL: http://people. maths. ox. ac. uk/richardsonm/SignalProcPCA. pdf (last access: 3.5. 2013). Aleš Hladnik Dr., Ass. Prof., Chair of Information and Graphic Arts Technology, Faculty of Natural Sciences and Engineering, University of Ljubljana, Slovenia ales. hladnik@ ntf. uni-lj. si*, 2009.

[6] A. M. Martínez and A. C. Kak, "Pca versus lda," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 23, no. 2, pp. 228–233, 2001.

[7] T. Kanungo, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman, and A. Y. Wu, "An efficient k-means clustering algorithm: Analysis and implementation," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 24, no. 7, pp. 881–892, 2002.

[8] M. Steinbach, L. Ertöz, and V. Kumar, "The challenges of clustering high dimensional data," in *New Directions in Statistical Physics*, pp. 273–309, Springer, 2004.

[9] R. Das, R. Dash, and B. Majhi, "Hyperspectral image classification based on quadratic fisher's discriminant analysis and multi-class support vector machine," *IETE Journal of Research*, vol. 60, no. 6, pp. 406–413, 2014.

[10] G. Camps-Valls, T. Bandos Marsheva, and D. Zhou, "Semi-supervised graph-based hyperspectral image classification," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 45, no. 10, pp. 3044–3054, 2007.

[11] A. Villa, J. A. Benediktsson, J. Chanussot, and C. Jutten, "Hyperspectral image classification with independent component discriminant analysis," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 49, no. 12, pp. 4865–4876, 2011.

[12] C. Shah, P. Watanachaturaporn, P. Varshney, and M. Arora, "Some recent results on hyperspectral image classification," in *Advances in Techniques for Analysis of Remotely Sensed Data, 2003 IEEE Workshop on*, pp. 346–353, IEEE, 2003.

[13] M. Rojas, I. Dópido, A. Plaza, and P. Gamba, "Comparison of support vector machine-based processing chains for hyperspectral image classification," in *SPIE Optical Engineering+ Applications*, pp. 78100B–78100B, International Society for Optics and Photonics, 2010.

[14] Q. Gu and J. Han, "Clustered support vector machines," in *proceedings of the sixteenth international conference on artificial intelligence and statistics*, pp. 307–315, 2013.

[15] Y. Yao, Y. Liu, Y. Yu, H. Xu, W. Lv, Z. Li, and X. Chen, "K-svm: An effective svm algorithm based on k-means clustering," *Journal of computers*, vol. 8, no. 10, pp. 2632–2639, 2013.